

IN THE UNITED STATES PATENT AND TRADEMARK OFFICE

#4
Jc872 U.S. PTO
10/058839
01/28/02

In Re the Application of : **Daisuke SHINOMIYA**
Filed: : **Concurrently herewith**
For: : **DYNAMIC LOAD SHARING SYSTEM....**
Serial No. : **Concurrently herewith**

Assistant Commissioner for Patents
Washington, D.C. 20231

January 28, 2002

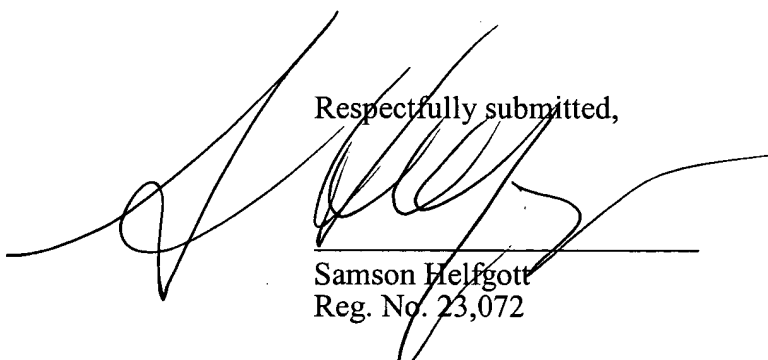
PRIORITY CLAIM AND SUBMISSION
OF PRIORITY DOCUMENT

S I R:

Applicant hereby claims priority under 35 USC 119 from **JAPANESE** patent application no. **2001-206856** filed **July 6, 2001**, a certified copy of which is enclosed.

Any fee, due as a result of this paper, not covered by an enclosed check, may be charged to Deposit Acct. No. 50-1290.

Respectfully submitted,



Samson Helfgott
Reg. No. 23,072

ROSENMAN & COLIN, LLP
575 MADISON AVENUE
IP Department
NEW YORK, NEW YORK 10022-2584
DOCKET NO.: FUJH 19.358
TELEPHONE: (212) 940-8800

日 本 国 特 許 庁
JAPAN PATENT OFFICE

1c872 U.S. P.
10/058839
01/28/02

別紙添付の書類に記載されている事項は下記の出願書類に記載されている事項と同一であることを証明する。

This is to certify that the annexed is a true copy of the following application as filed with this Office

出 願 年 月 日

Date of Application:

2001年 7月 6日

出 願 番 号

Application Number:

特願2001-206856

出 願 人

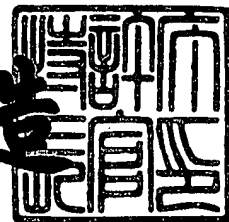
Applicant(s):

富士通株式会社

2001年10月 4日

特許庁長官
Commissioner,
Japan Patent Office

及川耕造



出証番号 出証特2001-3090789

【書類名】 特許願

【整理番号】 0150194

【提出日】 平成13年 7月 6日

【あて先】 特許庁長官 及川 耕造 殿

【国際特許分類】 H04L 12/46
H04L 12/48

【発明の名称】 仮想ルータを利用した動的な負荷分散システム

【請求項の数】 5

【発明者】

【住所又は居所】 神奈川県川崎市中原区上小田中4丁目1番1号 富士通株式会社内

【氏名】 四ノ宮 大輔

【特許出願人】

【識別番号】 000005223

【氏名又は名称】 富士通株式会社

【代理人】

【識別番号】 100094514

【弁理士】

【氏名又は名称】 林 恒▲徳▼

【代理人】

【識別番号】 100094525

【弁理士】

【氏名又は名称】 土井 健二

【手数料の表示】

【予納台帳番号】 030708

【納付金額】 21,000円

【提出物件の目録】

【物件名】 明細書 1

【物件名】 図面 1

特2001-206856

【物件名】 要約書 1

【包括委任状番号】 9704944

【ブルーフの要否】 要

【書類名】 明細書

【発明の名称】 仮想ルータを利用した動的な負荷分散システム

【特許請求の範囲】

【請求項 1】

1つの共通のアドレスを有する仮想ルータを構成する複数のルータ機能を有する装置と、

前記仮想ルータを通して、ネットワークに接続されるエンドシステムを有し、

前記仮想ルータを構成する複数のルータ機能を有する装置のうちのルータ機能を有する装置がマスタールータに設定され、他のルータ機能を有する装置がバックアップルータに設定され、

前記設定されたマスタールータは、ルーティング対象となるパケットの条件を動的に設定し、該設定されたルーティング対象となるパケットの条件を前記バックアップルータに通知し、前記複数のルータ機能を有する装置により前記ネットワークとエンドシステム間のルーティング処理を行う

ことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【請求項 2】 請求項 1 において、

前記バックアップルータは、前記マスタールータからルーティング対象となるパケットの条件を通知された際に、前記マスタールータに応答メッセージを返すことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【請求項 3】 請求項 1 において、

前記マスタールータは、前記バックアップルータにルーティング対象となるパケットの条件を通知した後に、前記バックアップルータに割当てたルーティング処理の対象となるパケットの条件を前記マスタールータ自身のルーティング処理の対象となるパケットの条件から外すことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【請求項 4】 請求項 2 において、

前記バックアップルータは、前記マスタールータに応答メッセージを返した後、一定期間ルーティング処理を行わない様に設定されることを特徴とする仮想ルータを利用した動的な負荷分散システム。

【請求項 5】請求項 2 において、

前記マスターータは、前記バックアップルータからの応答メッセージパケットを受けた時点で、割当てたルーティング対象となるパケットの条件をルーティング対象から外し、前記マスターータにおけるルーティング処理が終ったパケットのシーケンス番号を該当するバックアップルータに通知することを特徴とする仮想ルータを利用した動的な負荷分散システム。

【発明の詳細な説明】

【0 0 0 1】

【産業上の利用分野】

本発明は、データ通信を行うためのネットワーク中継装置（ルータ）、特に仮想ルータを利用した動的な負荷分散システムに関する。

【0 0 0 2】

【従来の技術】

近年はパーソナルコンピュータの急速な普及と共に、企業 I P 網（イントラネット）も普及し、多くの企業で利用されている。さらにコンピュータやネットワーク自体の高機能化、高性能化も進んでいる。

【0 0 0 3】

イントラネットは電子メールや WWW だけでなく、ストリーミングといったマルチメディアデータの通信にも用いられている。これらの通信用のプロトコルとして I P（インターネット・プロトコル）の占める割合が多くなってきており、この傾向は更に進んできている。

【0 0 0 4】

また、ネットワーク上で扱うデータも多様化しており、例えば、動画や業務用データのファイル転送など、伝送の遅れが生じると問題があるようなデータも同一ネットワーク上で取り扱われるようになっている。

【0 0 0 5】

通常、ネットワークはルータやレイヤ 3 スイッチを通して別のネットワークに接続されている。すなわち、自ネットワークとは異なるネットワークとの接続にルータが使われる。

【 0 0 0 6 】

このような状況においてルータに障害が発生した場合、別のネットワークのサーバ等と通信が行えなくなるため、業務が停止するなど深刻な問題となる。これを回避するためには、ルータを複数台設置すれば回避できるが、動的な経路切替を持たないパーソナルコンピュータ（PC）などのエンドシステムではデフォルト（初期設定）のゲートウェイを静的に指定しているため、障害時に対応することができない。

【 0 0 0 7 】

そこで、複数のルータで構成された仮想ルータを設定し、障害発生時にルータを切り替える技術を用いて耐障害性を上げることが提案され、実行されている。かかる技術としてインターネットの国際組織である I E T F（The Internet Engineering Task Force）で標準化された V R R P（Virtual Router Redundant Protocol）やホットスタンバイと呼ばれる技術がある。

【 0 0 0 8 】

ここで上記 V R R P は、ネットワーク上に存在する複数のルータで構成するバーチャル・ルータ（仮想ルータ：ネットワーク全体を1つのルータとみなす概念）の中から、動的に1つのルータをデフォルト・ルータ（デフォルト・ゲートウェイ）として選択するためのプロトコルである。

【 0 0 0 9 】

上記仮想ルータは複数のルータから構成されており、実際にルーティングを行うマスタールータとマスタールータに障害が発生した際に処理を引き継いで行うバックアップルータとに分かれて動作する。マスタールータに障害が発生すると、バックアップルータは自分がマスタールータとして動作し、通信を継続し、これにより信頼性の向上を実現する。

【 0 0 1 0 】

【発明が解決しようとする課題】

ここで、上記の技術は、通常動作しているルータ（マスタールータ）に障害が発生した場合や、監視区域内の経路に障害が発生した場合等に、待機系のルータに切り替える冗長構成の動作で信頼性を高める技術である。

【0011】

かかる場合、待機系のルータはマスタールータの障害有無を確認するためのパケットを送受信するために動作状態にあるが、ルーティングを行わないために装置のルーティング機能は生かされていないことになる。

【0012】

これに対応するため I E T F で標準化されている V R R P などでは、複数のグループを構成することで、静的に負荷分散を行うことができる。しかし、この方法は負荷の状態などを考慮しないため、特定のルータのみに負荷がかかってしまい、負荷分散にならないことがある。

【0013】

したがって、本発明の目的は、上記の問題点を解決し、動的に負荷分散を実現する仮想ルータを利用した動的な負荷分散システムを提供することにある。

【0014】

【課題を解決するための手段】

本発明では、上記の課題を解決するために仮想ルータを構成する複数のルータに、それぞれのルータでルーティングするパケットの条件を動的に設定し、同時に複数台のルータを動作させることで問題を解決する。

【0015】

また、ルーティングするパケットの条件を流量などの情報を元に定期的に見直しを行い、複数のルータが同程度の負荷で動作するようにすることで、ルーティング処理の負荷分散を効率よく行うことができる。

【0016】

仮想ルータを構成する場合、通常、マスタールータと呼ばれる実際にルーティング処理を行うルータと、マスタールータに障害が発生した時に代わりに動作するバックアップルータに分けられる。

【0017】

本発明は、通常、ルーティングを行わないで破棄するバックアップルータにルーティング処理を行うパケットの条件を設定し、それ以外のパケットについては従来通り破棄するように設定することで同時に複数台のルータを動かす。また、

グループを分けて静的に負荷分散を行うように設定しないため、エンドシステムで設定するデフをルートゲートウェイが1つでよい。

【0018】

上記課題を達成する本発明の仮想ルータを利用した動的な負荷分散システムの特徴は、1つの共通のアドレスを有する仮想ルータを構成する複数のルータ機能を有する装置と、前記仮想ルータを通して、ネットワークに接続されるエンドシステムを有し、前記仮想ルータを構成する複数のルータ機能を有する装置のうちのルータ機能を有する装置がマスタールータに設定され、他のルータ機能を有する装置がバックアップルータに設定され、前記設定されたマスタールータは、ルーティング対象となるパケットの条件を動的に設定し、該設定されたルーティング対象となるパケットの条件を前記バックアップルータに通知し、前記複数のルータ機能を有する装置により前記ネットワークとエンドシステム間のルーティング処理を行う。

【0019】

さらに、上記課題を達成する本発明の好ましい一態様として、前記バックアップルータは、前記マスタールータからルーティング対象となるパケットの条件を通知された際に、前記マスタールータに応答メッセージを返すことを特徴とする。

【0020】

また、上記課題を達成する本発明の好ましい一態様として、前記マスタールータは、前記バックアップルータにルーティング対象となるパケットの条件を通知した後に、前記バックアップルータに割当てたルーティング処理の対象となるパケットの条件を前記マスタールータ自身のルーティング処理の対象となるパケットの条件から外すことを特徴とする。

【0021】

さらにまた、上記課題を達成する本発明の好ましい一態様として、前記バックアップルータは、前記マスタールータに応答メッセージを返した後、一定期間ルーティング処理を行わない様に設定されることを特徴とする。

【0022】

さらに、上記課題を達成する本発明の好ましい一態様として、前記マスター

タは、前記バックアップルータからの応答メッセージパケットを受けた時点で、割当てたルーティング対象となるパケットの条件をルーティング対象から外し、前記マスタルータにおけるルーティング処理が終ったパケットのシーケンス番号を該当するバックアップルータに通知することを特徴とする。

【 0 0 2 3 】

本発明の特徴は、以下に図面を参照して説明する発明の実施の形態から更に明らかになる。

【 0 0 2 4 】

【発明の実施の形態】

以下に示す本発明の実施の形態例では、標準化された技術である VRRP に適用した場合に沿って説明するが、本発明の適用はこれには限定されない。すなわち、本発明は、各ベンダが独自の方式で実装している類似技術にも適用することが可能である。

【 0 0 2 5 】

ここで、本発明の実施の形態例を説明するにあたって、本発明の理解を容易とするべく VRRP の概要について先に説明する。

【 0 0 2 6 】

図 1 は、VRRP の概要を示す図である。VRRP は、図 1 に示されるようにパーソナルコンピュータ等の複数のエンドシステム 1-1 ~ 1-4 が HUB（あるいはスイッチ HUB）2 を通して仮想ルータ 3 に接続される。この仮想ルータ 3 は、複数のルータ（A）3-1，（B）3-2 で構成され、他のネットワーク NW に接続される。そして、障害発生時に動作するルータを例えば、ルータ 3-1 からルータ 3-2 に瞬時に切り替えることにより耐障害性を持たせることが可能である。

【 0 0 2 7 】

しかし、VRRP によらないシステムでは、複数のエンドシステム 1-1 ~ 1-4 に対し静的な経路設定しか行えないために、ルータ 3-1 あるいはルータ 3-2 をデフォルトゲートウェイとして直接指定した場合、障害発生時に他方のルータが正常動作をする場合であっても切り替えることが出来ない。

【0028】

これに対し、VRRPによるシステムでは、各ルータ3-1、3-2にはそれぞれ優先度が設定されており、仮想ルータ3のIPアドレスと同じアドレスを実アドレスとして持つルータが最も優先度が高く、実際にルーティング処理を行うマスタールータとして動作する。今、例としてルータ3-1の優先度が高く、これをデフォルトのマスタールータとする。

【0029】

マスタールータとなったルータ3-1は、仮想ルータ3を構成する自分以外のすべてのルータ（バックアップルータ）に対して、通知パケットを送信して、マスタールータであることを通知する。一方、マスタールータになれなかったルータ3-2はバックアップルータとしてマスタールータ3-1が動作しているかを確認するため、通知パケットの監視を行う。

一定期間、通知パケットが送信されてこなかった場合は、バックアップルータ3-2は、自分がマスタールータとして動作をする。ただし、バックアップルータが複数ある場合には、優先度の最も高いルータがマスタールータとして動作を継続し、それ以外のものはバックアップルータとして再びマスタールータの監視を行う。

【0030】

図2は、仮想ルータ3を構成する複数のルータのそれぞれの処理機能を説明するブロック図である。この図により各ルータの詳細機能をルータ3-1を代表例として説明する。

【0031】

図2における各機能ブロックは、ハードウェア若しくはソフトウェアで実現される。本例では、2組の送/受信ポート30、31（30-1、31-1）のうち、片側の送/受信ポート30、31だけにVRRPを使用している場合について説明する。

【0032】

各ルータは、VRRPデータテーブル32に、自ルータが属する仮想ルータ3についての情報を保持している。ここで保持されている情報とは、仮想ルータ3を構成するルータ3-1、3-2のIPアドレス、仮想ルータ3のIPアドレス

などである。

【0033】

初期状態において、VRRP処理部33でVRRPデータテーブル32の情報を基に、自ルータのIPアドレスと仮想ルータ3のIPアドレスを比較して自ルータがマスタールータとなるルータであるか否かを確認する。

【0034】

マスタールータとなった場合には、VRRPデータテーブル32の情報を基に通知パケットを構成し、送信ポート31から仮想ルータ3内の全てのバックアップルータに送信する。

【0035】

図3は、上記通知パケットのフォーマットであり、フォーマット内の情報の定義は次の様である。

- ・Version: VRRPのバージョンであり、現時点では1である。
- ・Type: VRRPのパケットタイプを示し、現時点では、1:通知パケットのみである。
- ・VRID: バーチャル・ルータIDであり、このIDによって同一仮想グループ内のルータを識別する。
- ・Priority: この値が大きいほど高いプライオリティとなる。
- ・Count IP Address: 仮想ルータに属するルータ数
- ・Auth Type: 認証方式を指定する。
- ・Adver Int: 通知パケットの送信間隔
- ・Check sum: パケットの誤り検査(チェックサム)
- ・IP Address: 仮想ルータを構成するルータのIPアドレスであって、構成するルータの数分ある。

【0036】

図2において、マスタールータとして動作を始めた瞬間から、図3に示すフォーマットの通知パケットを定期的を送信するためにタイマー34を計時すべく、VRRP処理部33からタイマー34に指示を送る。

【0037】

バックアップルータとなったルータ 3-2 は、受信ポート 30 からパケットを受信する。受信したパケットが VRRP 用の通知パケットかどうかをパケット判別部 35 で判断する。この判断は、図 8 A (IP ヘッダ) のプロコトルと上記図 3 のフォーマットの "Type" に基づいて判断する。

【0038】

バックアップルータでは、パケット判別部 35 でルーティングの対象となる通常のパケット (Type = 1 以外) であると判断する場合は、全て受信パケットを破棄する。バックアップルータは、ルーティング機能を有しないからである。受信パケットが VRRP 用の通知パケットである場合は、VRRP 処理部 33 に送り、処理の対象とする。

【0039】

VRRP 処理部 33 では、通知パケットを受け取った場合、タイマー 34 に対し、タイマー計数値を一旦リセットするように指示を送る。バックアップルータにおいて、このタイマー 34 の計数値は、マスタールータからの通知パケットの送信間隔を測るのに使われる。一定期間、パケットが送信されなければ、VRRP 処理部 33 により、バックアップルータからマスタールータへ切り替わるための準備を行う。

【0040】

かかる VRRP におけるルータでの動作により、障害時にルータを切り替えて処理を継続することができる。上記の VRRP による処理を前提とする本発明の実施例を以下に説明する。

【0041】

なお、以下の実施例においては、複数の仮想グループに設定されていないものとして説明をするが、複数の仮想グループを設定した場合でも、本発明を適用することが可能である。

【0042】

図 4 は、本発明を適用するルータの構成例ブロック図を示す。図 5 は、図 4 における分散処理用仮想ルータ制御部 (VR-CONT) 33 の詳細構成例ブロック図であり、図 6 は、更にパケット監視部 (PCKT-MON) 39 の詳細構成

例ブロック図である。

【0043】

上記したように、ルータの切替が発生するのは、マスタールータからの通知パケットが一定期間来なくなった場合を契機とする。そして、各バックアップルータは、そのタイミングでマスタールータに切り替わる。

【0044】

マスタールータに切り替わる場合には、その時点で、バックアップルータ群の中の最も優先度が高いルータが選ばれるため、マスタールータ以外のバックアップルータのうちいずれかに障害が発生したとしても、動作上問題がないことになる。

【0045】

しかし、本発明の目的はバックアップルータにおいても、ルーティング処理を行わすために全てのルータについて動作状況の監視を行うことが必要である。障害が発生した場合には他のルータが障害の発生したルータの代わりにルーティング処理を行わなければならない。

【0046】

したがって、本発明を適用する場合、「バックアップルータの故障の検出ができない」という問題点を解決することが必要である。このために、本発明では以下の様に機能を拡張する。

【0047】

すなわち、通知パケットを受け取ったバックアップルータは現在の自ルータの状況を通知するために、図3のフォーマット（但し、“Type”値を2とする）により応答を返す。これに対し、マスタールータは、一定期間内にバックアップルータから応答を返ってこない時点で、対象のルータに障害が発生していると判断する。そして、分散処理用データテーブル32に、当該ルータの状態を故障状態として登録する。

【0048】

また、本発明を実現する上で、送受される制御用のパケットは、実施例として、図3に示すVRRPの通知パケットを拡張する。各ルータに一斉に通知する必要のあるものは標準のVRRPと同様にマルチキャスト送信を利用する。また、

個別に送るパケットについては、登録されている各ルータの I P アドレス（図 3 参照）を元に宛先を指定して送信する。

【 0 0 4 9 】

以下かかる構成に基づき、本発明の処理動作を説明する。なお、ネットワーク構成を図 7 に示すように、図 1 の構成に対し更なるエンドシステム 1 - 5, 1 - 6 を、ルータ 4 及び、HUB 5 を通して接続する別のサブネットを有するものとして以下に説明する。

【 0 0 5 0 】

[初期動作]

図 7 のネットワークシステムにおいて、仮想ルータ 3 を構成する複数のルータ 3 - 1、3 - 2 は VRRP の機能に従って、マスタールータとバックアップルータのどちらかで動作を開始する。

【 0 0 5 1 】

図 7 で優先度の一番高いルータ（A）3 - 1 がマスタールータとなり、それ以外のルータ（B）3 - 2 がバックアップルータとなる。マスタールータとなったルータ 3 - 1 は、図 3 に示す通知メッセージにより分散処理用データテーブル 3 2 の内容をバックアップルータ 3 - 2 に送り、マスタールータとバックアップルータ間でデータを共有する。

【 0 0 5 2 】

これに対し、バックアップルータ 3 - 2 は、通知メッセージの”TYPE”を”2”に設定して、応答確認を返送する。これによりマスタールータでは先に説明したように、バックアップルータが正常であったか否かが判断出来る。

【 0 0 5 3 】

マスタールータ 3 - 1 は、更にルータとドメインを同じにするネットを意味するサブネットの構成情報としてサブネット内にいずれのエンドシステムが存在するかを調べるために情報要求メッセージをエンドシステム 1 - 1 ~ 1 - 4、1 - 5、1 - 6 に対して送信する（処理工程 P 3）。

【 0 0 5 4 】

上記のサブネットの構成情報を得るための情報要求メッセージとして、ICM

P (Internet Control Message Protocol) メッセージが利用される。ICMP メッセージの情報要求/応答メッセージは、標準勧告文書 RFC (Request For Comments) 792 で規定されている ICMP メッセージの一つであり、本来、自己のネットワークアドレスがわからない場合に使用される。

【0055】

ICMP メッセージのフォーマットを図 8 に示す。図 8 A はヘッダ部 (IP ヘッダ) であり、これに図 8 B の ICMP 情報要求/応答メッセージが付加される。図 8 A のプロトコルにより、ICMP 情報要求/応答メッセージを識別する。図 8 B の ICMP 情報要求/応答メッセージフォーマットにおいて、タイプは数字 15 で要求を、16 で応答を表す。また、識別子と順番番号は要求/応答比較用に使用する。

【0056】

かかる ICMP メッセージにおける送信元 IP アドレスのネットワーク部を "0" に設定し、サブネット内の全ての装置に対してブロードキャスト送信で情報要求メッセージを送信する。受信側は、送信元の IP アドレス (図 8 A) を宛先にして情報応答メッセージを返信する。

【0057】

返信された情報応答メッセージを受け取ったマスタールータ 3-1 は送信元アドレスを読み取り、サブネットの構成情報としてエンドシステムの IP アドレスを分散処理用データベース 32 に登録する。

【0058】

さらに詳細に説明すると、先ずマスタールータとなった時点で当該マスタールータ 3-1 の分散処理用仮想ルータ制御部 33 の処理部 331 (図 5 参照) から、通常の ICMP メッセージを処理するための機能ブロックである ICMP メッセージ処理部 37 に対して、ICMP 情報要求メッセージを送信するように指示を出す。

【0059】

これに対して各エンドシステムからの応答メッセージを受信ポート 30 により受け取ったマスタールータ 3-1 は、パケット判別部 35 において、ICMP メッ

セージであることを判断する（図8Aのプロコトルから判断出来る）。

【0060】

そして、ICMPメッセージであると判断すると、ICMPメッセージ処理部37に当該メッセージを転送する。

【0061】

そして、図8BのタイプからICMP情報応答メッセージであることを判断すると、分散処理用仮想ルータ制御部33の処理部331は、前記ICMPメッセージのヘッダ（図8A）から送信元のIPアドレスを取り出し、分散処理用データテーブル32に書き込む。

【0062】

分散処理用データテーブル32はさまざまな情報を保持しており、これらの情報は、大別すると3つの情報に分けられる。

【0063】

図9は、分散処理用データテーブル32に保持される情報の一例を示したものである。分散処理用データテーブル32は、仮想ルータ情報（図9A）、サブネット構成情報（図9B）及びパケット情報（図9C）の各情報テーブルを有している。

【0064】

前記の各エンドシステムからの応答メッセージに含まれる対応するIPアドレスは、サブネット構成情報（図9B）に記録される。

【0065】

ここで、ネットワークシステムとして、図7に示すように複数のエンドシステム1-1～1-4の他に、他のサブネット（HUM5を通してエンドシステム1-5、1-6を有する）と繋がるルータ（C）4がある場合は、当該ルータ4についてもサブネット構成情報テーブル（図9B）に登録される。図9Bでノード1、ノード2、ノード3・・・は、各サブネット構成におけるルータである。

【0066】

なを、ここではICMP情報要求メッセージを使ってエンドシステム1-1～1-4、1-5～1-6の情報を集める方法について示したが、他の方法も可能

である。一例として、サブネットのネットワークアドレスを基に配下のアドレスが全て存在するものとして各ルータに対して割当てを行う。実際にルータにパケットが到来した時点で、当該エンドシステムが存在することを確認し、サブネット情報として登録を行う。以降は、その情報を使って割り当てを行う方法である。

【0067】

上記の処理で、一連の動作サブネット内の構成を分散処理用データテーブル32に記録したら、この情報を基に各ルータに対していずれのパケットをルーティングするかを割り当てを行う。これに対し、該当のルータから確認メッセージが送られる。

【0068】

分散処理用データテーブル32に記録されている仮想ルータ情報(図9A)およびサブネット構成情報(図9B)を基に、いずれのパケットをどのルータでルーティングするかを決め、割り当てる。以降、各条件に対して割り当てられたルータを担当ルータと呼ぶ。

【0069】

分散処理用仮想ルータ制御部33の処理部331では、サブネット構成情報を取得した段階で割り当て処理部334に各パケットの条件に対して担当ルータを決める様に指示を送る。

【0070】

ここで、初期状態ではパケットの流量は不明であり、割り当て処理部334において初期動作としてパケットの割り当てを行う場合、仮想ルータを構成するルータの数及びサブネット構成情報しか有しない。このために単純にエンドシステムの総台数を仮想ルータを構成するルータの台数で割った数ずつ各ルータに割り当てる方法をとる。

【0071】

すなわち、単純にエンドシステムの総台数を、仮想ルータを構成するルータの台数で割った数ずつ各ルータに割り当てる。図7の例では、仮想ルータを構成するルータの台数は2、エンドシステムの総台数は6であるので、3ずつ各ルータ3

- 1, 3 - 2 に割当てることになる。

【 0 0 7 2 】

割当て処理部 3 3 4 では、データ管理部 3 3 2 を通じて、これらの情報を参照する。また、この際ルータ状態が故障のものには割当てを行わない。

【 0 0 7 3 】

担当ルータが決まった時点で、分散処理用仮想ルータ制御部 3 3 のメッセージ生成部 3 3 3 で割当てパケットを生成し、送信ポート 3 1 を通じて各ルータに対して通知を行う。かかる通知は、マルチキャストで一度に通知される。

【 0 0 7 4 】

図 1 0 は、図 3 の通知用フレームフォーマットを、本発明に適用するべく拡張して構成される拡張フレームフォーマットの例である。図 1 0 A は、上記の割り当てパケット用のフォーマットである。この時の T y p e 値は、" 3 " である。

【 0 0 7 5 】

図 1 0 B は、バックアップルータからの割り当て確認パケットである。この時の T y p e 値は、" 4 " である。

【 0 0 7 6 】

バックアップルータでは、上記の図 1 0 A の割当てパケットを受信すると、稼働の状態となり、通知された条件のパケットのルーティングを行うようになる。

【 0 0 7 7 】

割当てパケットを受信したバックアップルータは、分散処理用仮想ルータ制御部 3 3 (図 4) の T y p e 値判別部 3 3 0 (図 5) で割当てパケットであるか否かを判断する。

【 0 0 7 8 】

そして、分散処理用仮想ルータ制御部 3 3 の処理部 3 3 1 でルーティングを行うための処理を行う。そのため処理部 3 3 1 は、割当ての情報をデータ管理部 3 3 2 を通じて分散処理用データテーブル 3 2 のパケット情報 (図 9 C) に書き込むと同時に判別用データテーブル 4 0 (図 4 参照) にパケットの情報を書き込む。

【 0 0 7 9 】

判別用データテーブル40は、各条件とそれをどのように扱うかの状況について記述されたテーブルであり、その記述されたデータの一例が図13に示される。各条件に対して該当するパケットをルーティングするのか、破棄するのかについて記述されている。

【0080】

また、分散処理用仮想ルータ制御部33のメッセージ生成部333に割当て確認パケット（図10B）を生成するように指示を出し、マスタールータ3-1に対して割当て確認パケットを返信する。

【0081】

割当て確認パケットを受信したマスタールータ3-1は、該当するパケットをルーティング対象から外し、破棄するための処理を行う。また、このとき分散処理用データテーブル32の該当ルータのルータ状態を休止から稼動に切り替える。

【0082】

受信したパケットが割当てパケットであるとType値判別部330で判断されると、処理部331ではデータ管理部332に対して分散処理用データテーブル32の該当ルータのルータ状態を休止から稼動へ変更するように指示を出す。

【0083】

また、そのときマスタールータ3-1の判別用データテーブル40に対して、該当するパケットのルーティングを禁止するように指示を出す。

[パケット複製の回避]

マスタールータ3-1から各バックアップルータ3-2へ割当てパケットを送信することで、前述のように各ルータがルーティング処理を開始するが、マスタールータ3-1の判別用データテーブル40から該当するパケットをルーティングの対象から外すまでの期間、パケットの複製（マスタールータとバックアップルータによる同時のルーティング）が発生する恐れがある。

【0084】

本発明では2つの方法でこの問題を解決する。どちらの方法によって解決するかは仮想ルータ情報の動作モード（図9A）として記録されている。これは各ルータに手動で設定され、マスタールータ3-1に設定されている動作モードで動作

する。したがって、マスタータ3-2が障害により切り替わった場合には、新しくマスタータになったルータに設定されている動作モードで動作する。

【0085】

[タイマーによる回避]

各バックアップルータ3-2はマスタータ3-1からの割当てパケット(図10A)を受信した時点で、ルーティングを開始する処理を一定期間遅らせる。この間に、割当て確認パケット(図10B)がマスタータ3-1で受信される。

【0086】

これにより該当パケットがマスタータ3-2によるルーティングの対象から外れるため、パケットの複製が発生しない。具体的な処理は次のようになる。

【0087】

各バックアップルータ3-2がマスタータ3-1からの割当てパケット(図10A)を受けた時、分散処理用仮想ルータ制御部33の処理部331はデータ管理部332を通じて分散処理用データテーブル32に割当ての情報を記録するように指示を出す。

【0088】

同時に、メッセージ生成部333に対して割当て確認パケットをマスタータ3-1宛てに送信するよう指示を出す。このとき、タイマー34に対して一定期間計数を遅らすようにタイマー動作に対する指示を出す。これにより判別用データテーブル40に対して、該当パケットの情報を書き込むことを行わせない様にすることが出来る。

【0089】

タイマー34から一定期間の経過が通知された時点で、判別用データテーブル40に対して、該当パケットの情報を通知し、ルーティングの対象として処理させる。

【0090】

この方法は、次に示す方法よりも単純な処理で実現できるメリットがあるが、パケットロスの可能性があるという問題もある。

【0091】

[シーケンス番号を使った回避]

上記タイマーによる回避方法でパケットの複製を防ぐことは出来るが、パケットロスが発生する可能性がある。そこでパケットロスも合わせて防ぐための方法としてマスタータ3-1が、割当て確認パケットを受け、該当パケットをルーティングの対象から外した時点で、現在どのパケットまでルーティングが終了したかを調べる。

【0092】

そして、バッファ処理部38（図4参照）にバッファリングされる該当パケットのTCPシーケンス番号を使って該当のバックアップルータ3-2に対して通知する。TCPシーケンス番号は、図11Cのシーケンス番号通知パケットを使用して送られる。

【0093】

マスタータ3-1の処理部331で割当て確認パケット（図10B）が到達したことを確認すると、そのメッセージを返してきたバックアップルータ3-2に割当てた条件に対する、パケットの現在の状態を調べる。

【0094】

具体的には処理部331が、パケット監視部39の該当パケット識別部（図6，395）に対して、バッファ中の該当パケットの情報を返すように指示を送る。

【0095】

該当パケット識別部394は、バッファチェック部392に対して、バッファ処理部38の中に蓄えられているパケットの中から該当するパケットのうち、各条件毎に最も古いパケットを取り出す。

【0096】

バッファ処理部38に蓄えられているパケットに対し、バックアップルータ3-2に対して割当てたIPアドレスから送られてきたパケットであるかを古い順に調べる。パケットが該当パケットである場合で、かつTCP (Transmission Control Protocol:コネクションという接続関係を結ぶ形態のトランスポート層プ

ロトコル) のパケットであれば、そのパケットの送信元 IP アドレス、宛先 IP アドレス、ポート番号およびシーケンス番号をチェックし処理部 331 に対して通知する。

【0097】

また該当したパケットについては、その時点で破棄され、以降のバッファに蓄えられた同じ条件のパケットについては情報を通知することなくパケットの破棄だけを行う。

【0098】

もし、該当する IP アドレスからのパケットであって、UDP (User Datagram Protocol: コネクションレスというデータ転送するだけの形態のトランスポート層プロトコル) のパケットであれば、無条件で破棄を行う。

【0099】

また、この処理の結果バッファ処理部 38 の中に該当する IP アドレスからの TCP パケットが見つからなかった場合は、処理部 331 に対して該当なしの通知を行う。

【0100】

各条件毎の情報をもらった処理部 331 は、その情報をメッセージ生成部 33 に通知し、シーケンス番号通知パケットを生成し、該当のバックアップルータ 3-2 に対して送信するように指示を出す。

【0101】

シーケンス番号通知パケット (図 10A) を受けたバックアップルータ 3-2 はバッファ処理部 38 に蓄えられているパケットの中から該当する TCP シーケンス番号をもつパケットより前の同じ条件に該当するパケットを破棄し、通知されたシーケンス番号以降のパケットからルーティングを開始する。

【0102】

これにより、パケットの複製を防ぐと同時にパケットロスも防ぐことができる。

【0103】

シーケンス番号通知パケットを受けたバックアップルータ 3-2 の処理部 33

1 は、パケット監視部 39 の該当パケット識別部 394 に対して送られたシーケンス番号以前の該当するパケットを破棄するように指示を出す。

【0104】

該当パケット識別部 394 は、各条件毎の該当パケットの情報と一致するパケットを探すようにバッファチェック部 392 に対して指示を出す。バッファチェック部 392 は、バッファの先頭から該当するパケットがないかのチェックを行う。

【0105】

このチェックにおいて、条件は一致するが該当パケットの情報と一致しないものは削除していく。該当のパケットが出てきた時点でその条件のパケットのチェック処理を止める。

【0106】

この方法では、UDP のパケットが破棄されてしまう可能性があるが、UDP の場合は削除されるパケットが少なければあまり問題にならないためにここでは考慮しない。

【0107】

このとき、マスターータ 3-1 からのシーケンス情報通知パケット内で割当て条件に関してシーケンス番号などの情報が「該当なし」で通知されたときに、バックアップルータ 3-2 内のバッファに割当て条件のパケットが存在していた場合は、バックアップルータ 3-2 のバッファのどの位置にパケットがあるかによって判断する。バッファの後半にあれば該当パケットを残すが、前半にあった場合は削除する。

【0108】

この方式の場合、割当てパケットを受けた時点で、判別用データテーブル 40 に対象となるパケットの情報を書き込む。そしてバッファ処理部 38 に対して許可があるまで、転送機能部 36 に対してバッファに蓄えられたパケットを送らないように指示を出す。これはシーケンス番号通知パケットが来た時点で処理部 331 から指示が出され解除される。

【0109】

以上の動作を図14のフローに纏めて示す。

【0110】

図14において、パケット複製の回避を含めた初期動作における、マスタータ3-1、バックアップルータ3-2およびエンドシステム1-1~104, 1-5~1-6間でのパケットの流れを示す。

【0111】

二重線で示されているパケットの通知（処理工程P1）はマスタータ3-1とバックアップルータ3-2間の通知パケットの送受であって、標準のVRRP通知パケットと同様に定期的に行われる。

【0112】

処理工程P2では、マスタータ3-1から各エンドシステム1-1~1-4、1-5~1-6へ情報収集のためのパケットを送信し、それに対する応答パケットの流れである。

【0113】

エンドシステム1-1~1-4, 1-5~1-6の情報が集まった時点でマスタータ3-1で割当ての処理を行う（処理工程P3）。

【0114】

割当て処理（処理工程P3）が終ると、その結果をバックアップルータ3-2に対して通知し、その通知を受け取ったバックアップルータ3-2は、それに対して応答を返信する（（処理工程P4））。

【0115】

ついで、マスタータ3-1は、バックアップルータ3-2に対し、先に説明した回避方法として、シーケンス番号の通知によって、パケットの複製の回避を行う（処理工程P5）。シーケンス番号の通知は、図11Cに示すシーケンス番号通知パケットにより通知される。このシーケンス番号の通知による回避の場合は、タイミングT2でルーティングを開始する。また、別の回避方法としてのタイマー処理による場合は、タイミングT1でルーティング処理を開始する。

[割当ての見直しと再割当て]

ここで、ルーティングの割当ての結果、分散処理が行われるが、パケット量の

増減により偏りが発生する可能性がある。そこで必要に応じて見直しを行う。

【0116】

見直しは、マスタータ3-1から行う場合と、バックアップルータ3-2から行う場合とがある。両ケースの違いは、見直しが必要かどうかを判断するための情報を集めるのにマスタータ3-1が自発的に情報集めを行うか、バックアップルータ3-2がマスタータ3-2に対して要求を出してマスタータ3-2に処理させるかである。

【0117】

見直しが必要かどうかの判断のための情報が集まった後に、見直しが必要かどうかの判断や、その後の再割当て処理は上記いずれのケースの場合ともに同じ処理が行われる。

【0118】

以下に実施例として、バックアップルータ3-2から要求が出される場合の処理の流れを示す図15のシーケンス図に沿って、各機能ブロックでどのような処理がなされるかを説明する。

【0119】

通常、バックアップルータ3-2や、マスタータ3-1では、ルーティング処理の対象となるパケットの監視を行っている。パケットの監視はパケット監視部39で行っており、2つの情報を監視している。

【0120】

一つはバッファのあふれ回数であり、もう一つは各条件に該当するパケットの流量である。ここでいう流量はパケットの長さを考慮してbpsで表される。

【0121】

バッファのあふれ監視はバッファあふれ監視部390（図6参照）で行われる。バッファあふれ監視部390はバッファ処理部38においてバッファあふれが発生した時点でバッファあふれカウンタ部391に対してバッファあふれが起きたことを通知する。

【0122】

バッファあふれカウンタ部391では、単位時間あたりのバッファあふれ回数

をカウントしている。そのカウント値の情報は条件判別部 395 と流量記録制御部 396 に通知される。

【0123】

また各条件に該当するパケットの流量はバッファチェック部 392 によって監視される。バッファチェック部 392 では、バッファ処理部 38 を通り抜けてルーティング処理されるパケットの情報をチェックする機能も持っている。各パケットの情報は短い期間でチェックされ、流量チェック部 393 で条件毎に集計され、流量記録制御部 396 に送られる。

【0124】

流量記録制御部 396 は分散処理用データテーブル 32 に対して一定間隔で、現時点での単位時間あたりのバッファあふれ回数と、各条件毎の過去の一定期間内での平均のパケット流量のデータ（単位はbps）を書き込む。

【0125】

一方、条件判別部 395 に送られた単位時間あたりのバッファあふれ回数と各条件毎の平均パケット流量は、条件判別部 395 で見直し要求を行う必要があるかどうかの判断用の情報として使われる。

【0126】

条件判別部 395 は、比較用の条件として閾値が手動で設定されている。設定された比較閾値は、パケット流量の情報やバッファあふれ回数の情報と比較される。

【0127】

条件判別部 395 はこれらの値が閾値よりも大きい場合、見直しの要求をする必要があると判断し、分散処理用仮想ルータ制御部 33 の見直し処理部 335 に対して見直しの確認を行うように要求を送る。

【0128】

バックアップルータ 3-2 の場合、見直し処理部 335 に見直しの確認要求がきたとき、見直し処理部 335 は、メッセージ生成部 333 に対して見直し要求パケット（図 11D）を生成し、マスタルータ 3-1 宛てに送信するように指示する。

【0129】

バックアップルータ3-2から見直し要求パケットを受け取ったマスタルータ3-1は、Type値判別部330で見直し要求であること（Type値が6である）を確認すると、見直し処理部335にその旨を通知する。

【0130】

見直し処理部335では、バックアップルータ3-2から見直し要求がきたので、実際に再割当てを行う必要があるか否かの判断情報を収集するために、メッセージ生成部333に対して流量通知要求パケット（図11E）をバックアップルータ3-2宛てに送信するように指示を出す。

【0131】

バックアップルータ3-2は、Type値判別部330で流量通知要求パケットを受信したこと（Type値が7である）を確認すると、流量通知パケットを見直し処理部335に対して通知する。見直し処理部335では、データ管理部332を通じて分散処理用データテーブル32に記録されている流量に関する情報（パケット情報として記録されている）を取り出す。

【0132】

取り出されたパケット情報は、メッセージ生成部333に対して、それらの情報を元に流量通知応答パケット（Type値が8である：図12F）を生成し、マスタルータ3-1に対して返信するように指示を出す。

【0133】

マスタルータ3-1では、流量通知応答パケットを受け取ると見直し処理部335に通知する。ついで、データ管理部332に対してそれらのデータをパケット情報として分散処理用データテーブル32に書き込む。このような手順で、判断に必要な情報を収集する。

【0134】

この見直しが必要かどうかの判断は、前記したようにバックアップルータ3-2からのトリガで行われる以外に、マスタルータ3-1自身から自分のパケット流量が多い場合に行われる。この場合、条件判別部395から見直し処理部335に通知が送られた時点で、見直し処理部335はメッセージ生成部333に対

して流量通知要求パケット（図11E）をバックアップルータ3-2宛てに送信するように指示を出す。

【0135】

これ以降の情報収集の流れはバックアップルータ3-2からの要求をトリガにして動作した場合と同じ流れである。最終的にバックアップルータ3-2からの流量通知応答パケット（図12F）を受け取ると、その情報を分散処理データテーブル32に書き込む。

【0136】

上記のように、マスタールータ3-1がパケットの流量を監視した結果、見直し確認が必要であると決定した場合、あるいはバックアップルータ3-2がパケットの流量を監視し、見直し確認が必要であると決定して、マスタールータ3-1に見直し確認を要求した場合にマスタールータ3-1は見直しの必要を確認する。

【0137】

図15のシーケンス図ではバックアップルータ3-2で見直し確認の必要が出てきたため、マスタールータ3-1に見直し確認要求をしている（処理工程P10）。要求を受け取ったマスタールータ3-1は、バックアップルータ3-2に対して、流量に関する情報を通知するように指示を出し、これに対しバックアップルータ3-2は当該要求に応じて情報を通知する（処理工程P12）。

【0138】

判断に必要な情報が集まると、マスタールータ3-1は見直しが必要かどうかの判断を行う（処理工程P13）。判断基準は見直し処理部335に手動で設定されており、各ルータに割当てられているパケットの平均流量の総量の差が設定された値を上回っている場合か、あるいは、単位時間当たりのバッファあふれ回数の差が設定された値を上回っていると判断された場合である。

【0139】

この判断基準に合う時、再割当ての処理を開始するために割当て処理部334に再割当てを行うように指示を送る。

【0140】

割当て処理は初期状態と同様にエンドシステム1-1～1-4、1-5～1-

6の構成を見直し、割当てアルゴリズムによって割当てを行う。割当ての方法としては、流量順に各条件を並べて同程度の総流量になるように各ルータに割振っていく方法などがある。

【0141】

このとき、新たに追加されているエンドシステムがあった場合は現在の流量を0のものとして割当てを行う。割当てが決まった時点で、初期動作の時と同様にバックアップルータ3-2に対して通知を行う。

【0142】

図15のシーケンス図の流れではP13のタイミングでマスタールータ3-1は、見直しが必要かどうかを判断し、必要な場合に再割当ての処理が行われる。

【0143】

図15で、処理工程P14～P16の処理は、先に示した通常の初期動作の場合と同じ手順（図14，処理工程P2～P4）である。

【0144】

また、これらの見直し確認の処理が頻繁に発生すると判断用の情報収集の処理などで各ルータに負担がかかる。したがって、バックアップルータ3-2であれば、見直し要求パケットをマスタールータ3-1に送信した場合に、またマスタールータ3-1であれば自発的に流量通知要求パケット（図11E）の送信を行った場合に、それぞれの処理を一定期間行わないようにする。

【0145】

このために、それぞれの見直し処理部335では、該当する流量通知要求パケットの生成をメッセージ生成部333に指示を出した時点で、タイマー34へ一定期間動作を行わないようにするためのタイマー停止の指示を送る。したがって、タイマー34からの解除通知が来るまで見直し処理部335では、該当する処理を行わない。すなわち、上記の割り当て処理の流れにおいて、再割当てが行われたか行われなかったにかかわらず、見直し処理部335で処理を行わないようにする。

[別装置による管理]

上記した本発明の動作において、ルータがルーティング以外の機能部分で処理

する負荷が増えることが考えられる、かかる場合は必ずしも望ましくない場合が想定される。

【0146】

これに対応するために、本発明において、マスターータ3-1で行われる処理を代わりに処理するサーバを別途設け、当該サーバに前記処理をさせることも可能である。

【0147】

図16は、かかる場合のマスターータ3-1で行う処理を代行する代替処理用の代行サーバ6を設けた場合のシステム構成を示す図である。この代行サーバ6は、分散処理用仮想ルータ3に関する処理とICMP (Internet Control Message Protocol) に関する処理を行う機能を持つ必要があり、仮想ルータ3を構成する特別なルータとして扱われる。

【0148】

実際に、代行サーバ6に処理を代替させる場合の動作の流れを図17のシーケンス図に沿って説明する。

【0149】

代行サーバ6は、仮想ルータ3を構成するルータの一つとして事前に登録されている必要がある。この時、代行サーバ6は、ルータ状態が代替処理用のサーバであることを示す識別番号で登録されている。

【0150】

マスターータ3-1は、通常のVRRPと同様に通知パケットを送信する。通知パケットに対する応答は代行サーバ6、バックアップルータ3-2共に通常と同じように行われる。二重線で記述されているパケットの流れ（処理工程P20, p21）は従来のVRRPの通知パケットと同じで定期的に送受信され、を互いに動作しているかどうかを確認する。

【0151】

マスターータ3-1から通知パケットを受け取り、仮想ルータ3が動作していることを確認した代行サーバ6は、エンドシステム1-1~1-4, 1-5~1-6に対して情報を要求するパケットを送信する（処理工程P22）。各エンド

システムは代行サーバ6に対して応答を返す。かかる動作は、図14に示したルータ3-1、3-2のみで処理される場合と同じである。

【0152】

各エンドシステムの情報が集まった時点で割当てを開始する。代行サーバ6はマスタールータ3-1で行う場合と同様に割当てを行う（処理工程P23）。

【0153】

割当てが決まった段階で、代行サーバ6はマスタールータ3-1とバックアップルータ3-2に対して割当てを通知する（処理工程P24）。この時、ルータ3-1は、自分がマスタールータであるため情報を分散処理用データテーブル32に記録を行い、バックアップルータ3-2のみから応答が戻り、マスタールータ3-1からは応答を返さない。

【0154】

代行サーバ6は、バックアップルータ3-2からの応答メッセージを受け取ると、マスタールータ3-1に対して転送する（処理工程P25）。このため、タイマー34を用いてパケットの複製を回避する方法を適用する場合、タイマーの時間を長めに設定する必要がある。この時の、バックアップルータ3-2でルーティングを開始するのは図17において、T2のタイミングである。

【0155】

また、シーケンス番号を使用する場合には、点線のパケットの送信がある（処理工程P26）。これは、マスタールータ3-1から直接行われ、代行サーバ6を使用しない場合の処理と全く同じ処理を行う。

【0156】

次に、図16の構成における代行サーバ6を用いる場合の見直しの際の動作の手順を示す図18を参照して、見直し動作を説明する。見直しは、マスタールータ3-1あるいはバックアップルータ3-2から要求が代行サーバ6宛てに送られる（処理工程P30）。

【0157】

見直し要求を受け取った代行サーバ6はマスタールータ3-1とバックアップルータ3-2に対して、情報の要求を行う。情報の要求メッセージを受け取ったマ

スタルータ 3-1 とバックアップルータ 3-2 は、現在のパケットの流量および、一定期間内でのバッファあふれ回数を代行サーバ 6 に回答する（処理工程 P 3 1）。

【0 1 5 8】

代行サーバ 6 は、これらの情報から見直しが必要か否かの判断を行う。判断基準や見直しの判断処理などはマスタルータ 3-1 が直接行う場合と同様である（処理工程 P 3 2）。

【0 1 5 9】

見直しが必要と判断した場合には、図 1 7 に示した初期動作と同様にエンドシステム 1-1 ～ 1-4、1-5 ～ 1-6 からの情報を集める（処理工程 P 3 3）。その後、再割当ての処理（処理工程 P 3 2）を行い、マスタルータ 3-1 とバックアップルータ 3-2 に対して通知を行う（処理工程 P 3 3）。

【0 1 6 0】

なを、代行サーバ 6 が故障してこれらの処理が行えなくなった場合、マスタルータ 3-2 は、代行サーバ 6 の状態を（代替処理用サーバ：故障）に設定し、サーバを使わないでマスタルータ 3-1 で処理を行う。

【0 1 6 1】

[本発明を利用したサービス]

本発明を使って、IP の仮想専用線サービスを提供するキャリアや ISP（インターネットサービスプロバイダ）のネットワークを介して、複数の事業所を接続することが考えられる。

【0 1 6 2】

説明のために、単純に 2 点間の場合の構成例を図 1 9 に示す。図 1 9 において、キャリアネットワーク CNW を通して企業における本店ネットワーク NW 1 と支店ネットワーク NW 2 が接続されている。

【0 1 6 3】

本店ネットワーク NW 1 とキャリアネットワーク CNW 間及び、支店ネットワーク NW 2 とキャリアネットワーク CNW 間に本発明による仮想ルータ 3 が備えられている。この仮想ルータ 3 に前記各実施の形態により説明した本発明の特徴

が適用可能である。

【 0 1 6 4 】

(付記 1)

1 つの共通のアドレスを有する仮想ルータを構成する複数のルータ機能を有する装置と、

前記仮想ルータを通して、ネットワークに接続されるエンドシステムを有し、

前記仮想ルータを構成する複数のルータ機能を有する装置のうちのルータ機能を有する装置がマスタールータに設定され、他のルータ機能を有する装置がバックアップルータに設定され、

前記設定されたマスタールータは、ルーティング対象となるパケットの条件を動的に設定し、該設定されたルーティング対象となるパケットの条件を前記バックアップルータに通知し、前記複数のルータ機能を有する装置により前記ネットワークとエンドシステム間のルーティング処理を行う

ことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 6 5 】

(付記 2) 付記 1 において、

前記バックアップルータは、前記マスタールータからルーティング対象となるパケットの条件を通知された際に、前記マスタールータに応答メッセージを返すことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 6 6 】

(付記 3) 付記 1 において、

前記マスタールータは、前記バックアップルータにルーティング対象となるパケットの条件を通知した後に、前記バックアップルータに割当てたルーティング処理の対象となるパケットの条件を前記マスタールータ自身のルーティング処理の対象となるパケットの条件から外すことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 6 7 】

(付記 4) 付記 2 において、

前記バックアップルータは、前記マスタールータに応答メッセージを返した後、

一定期間ルーティング処理を行わない様に設定されることを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 6 8 】

(付記 5) 付記 2 において、

前記マスタールータは、前記バックアップルータからの応答メッセージパケットを受けた時点で、割当てたルーティング対象となるパケットの条件をルーティング対象から外し、前記マスタールータにおけるルーティング処理が終ったパケットのシーケンス番号を該当するバックアップルータに通知することを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 6 9 】

(付記 6) 付記 5 において、

前記バックアップルータは、前記マスタールータから通知を受けたシーケンス番号を基に、バッファリングされたパケットのうち、既に前記マスタールータでルーティングされたパケットを破棄し、それ以降のパケットからルーティングを行うことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 0 】

(付記 7) 請求項 1 において、

前記バックアップルータは、自装置でルーティングするパケットの流量情報を監視する手段を有することを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 1 】

(付記 8) 付記 7 において、

前記バックアップルータは、前記監視手段で監視される流量情報により、パケット流量が所定値以上になる時、前記マスタールータに対して、ルーティング処理の対象となるパケットの条件の見直しを要求することを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 2 】

(付記 9) 付記 7 において、

前記マスタールータは、前記バックアップルータの監視手段で監視される流量情

報を収集し、パケットの流量が所定値以上になる時、ルーティング処理の対象となるパケットの条件の見直しを行うことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 3 】

(付記 1 0) 付記 9 において、

前記前記マスタールータにおけるバックアップルータからの流量情報の収集は、前記バックアップルータからの要求に応じて実行されることを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 4 】

(付記 1 1) 付記 9 において、

前記パケットの条件の見直しにより、パケットのルーティングを担当するルータの再割当てを行うことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 5 】

(付記 1 2) 付記 1 1 において、

前記パケットのルーティングを担当するルータの再割当ての有無にかかわらず、
前記バックアップルータからの見直し要求を一定期間禁止することを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 6 】

(付記 1 3) 付記 1 1 において、

前記パケットのルーティングを担当するルータの再割当ての有無にかかわらず、前記マスタールータは、情報収集を一定期間行わないことを特徴とする仮想ルータを利用した動的な負荷分散システム。

【 0 1 7 7 】

(付記 1 4) 付記 1 乃至 1 3 にいずれかにおいて、

更にサーバ装置を備え、

前記サーバ装置によりルーティングするパケット情報の変更やエンドシステムの構成情報の取得を行うことを特徴とする仮想ルータを利用した動的な負荷分散

システム。

【0178】

(付記15) 請求項1において、

前記ネットワークは、IPの仮想専用線サービス(IP-VPNサービス)を提供しているキャリアやISP(インターネットサービスプロバイダ)等のネットワークであることを特徴とする仮想ルータを利用した動的な負荷分散システム。

【0179】

【発明の効果】

以上図面に従い説明したように、本発明により、従来の仮想ルータによる冗長構成と比較して、複数のルータ3-1, 3-2を同時に動かすため、単純に処理可能なパケット量が増えるというメリットがある。また、冗長構成で静的な負荷分散を実現した場合に比べて、各端末からのパケット量を類推して静的に負荷分散の設定を行う手間がかからず、負荷の変化にも柔軟に対応できるというメリットがある。

【0180】

当然に、冗長構成をとっているため障害に対して高い信頼性を持つというメリットも維持出来る。

【0181】

これらのメリットにより、本発明を適用することで良好なレスポンスと高い信頼性および負荷分散を同時に実現するサービスを提供することができる。

【0182】

さらに、上記の特徴を有する本発明を利用することで、従来の仮想ルータによる冗長構成をした場合に比べて、複数のルータが同時に動くようになるため、従来よりも機器および経路の帯域を有効に使うことができる。また、デフをルトゲートウェイを複数用意して静的に負荷分散を行う場合に比べて、負荷の偏りに対して柔軟に対応できることから、負荷分散を効率よく行うことができる。

【0183】

本発明では、設定するデフォルトゲートウェイは各エンドシステムで同じ設定をすれば済む。したがって、静的に負荷分散を設定する場合のように、事前に各

エンドシステムのパケットの流量を細かく考慮する必要がない。これにより、ネットワーク設計を含めた設定が非常に楽に行える。

【 0 1 8 4 】

また、上記のメリットがあることから、IP-VPNサービスを提供しているキャリアやISPのネットワークにおいて、本発明の機能を持つルータを使うことで、良好なレスポンスと高い信頼性と負荷分散を実現したサービスを同時に提供することができる。

【図面の簡単な説明】

【図 1】

VR RPの機能を説明する概要を示す図である。

【図 2】

VR RP機能を実現するルータの機能ブロック図である。

【図 3】

VR RPで使われるメッセージのフォーマットである。

【図 4】

本発明の機能を実現するルータの機能ブロック図である。

【図 5】

図 4 の構成における分散処理用仮想ルータ制御部の詳細構成例ブロック図である。

【図 6】

図 4 の構成におけるパケット監視部 3 9 の詳細構成例ブロック図である。

【図 7】

別のサブネットを有するネットワーク構成を示す図である。

【図 8】

ICMPメッセージのフォーマットを示す図である。

【図 9】

分散処理用データテーブル 3 2 に保持される情報の一例を示したものである。

【図 1 0】

図 3 に対する本発明に適用される拡張フォーマットで、割り当てパケットと、

割り当て確認パケットである。

【図 1 1】

図 3 に対する本発明に適用される拡張フォーマットで、シーケンス通知パケット、見直し要求パケット及び、流量通知パケットである。

【図 1 2】

図 3 に対する本発明に適用される拡張フォーマットで、流量通知応答パケットである。

【図 1 3】

汎用データテーブルの一例である。

【図 1 4】

パケット複製の回避を含めた初期動作における、マスタールータ、バックアップルータおよびエンドシステム間でのパケットの流れを示す図である。

【図 1 5】

バックアップルータ 3 - 2 から要求が出される場合の処理の流れを示す図である。

【図 1 6】

マスタールータで行う処理を代行する代替処理用の代行サーバを設けた場合のシステム構成を示す図である。

【図 1 7】

代行サーバに処理を代替させる場合の動作の流れを示す図である。

【図 1 8】

代行サーバを用いる場合の見直しの際の動作手順を示す図である。

【図 1 9】

I P の仮想専用線サービスを提供するネットワークを介して、複数の事業所を接続する場合の本発明の適用を説明する図である。

【符号の説明】

1 - 1 ~ 1 - 6 エンドシステム

2、5 HUB

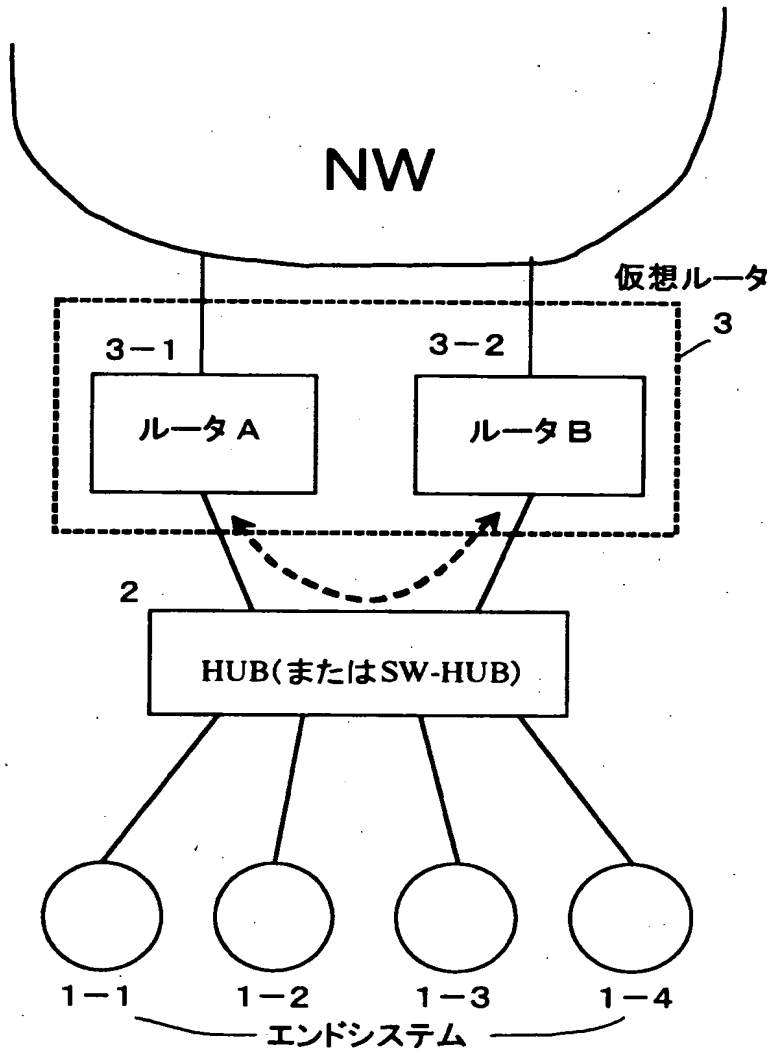
3 仮想ルータ

3 - 1 , 3 - 2 、 4 ルー タ

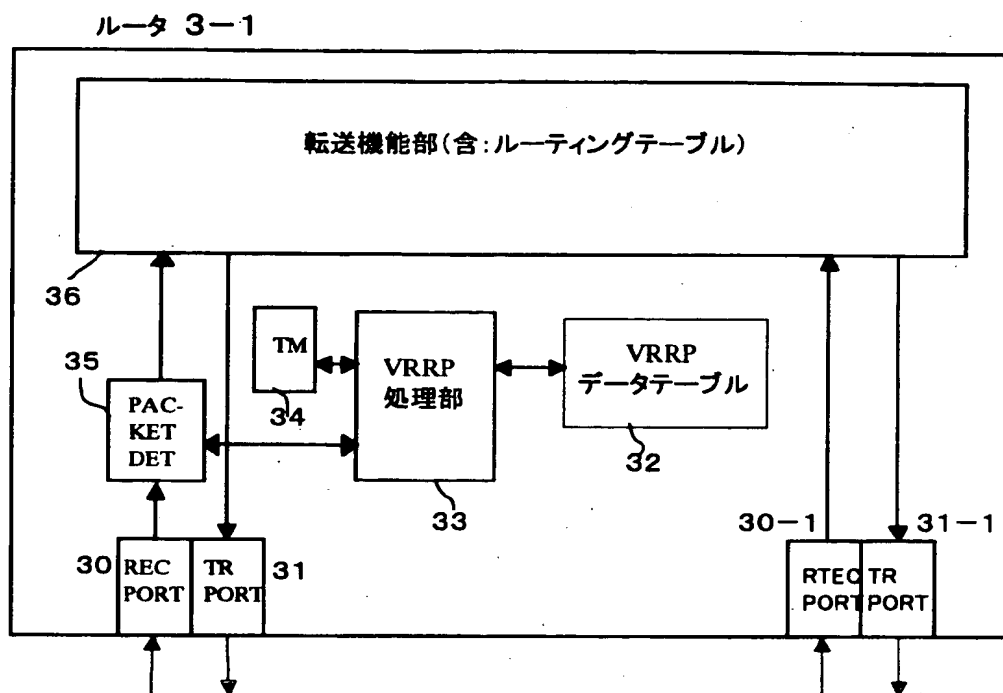
6 代行サーバ

【書類名】 図面

【図 1】



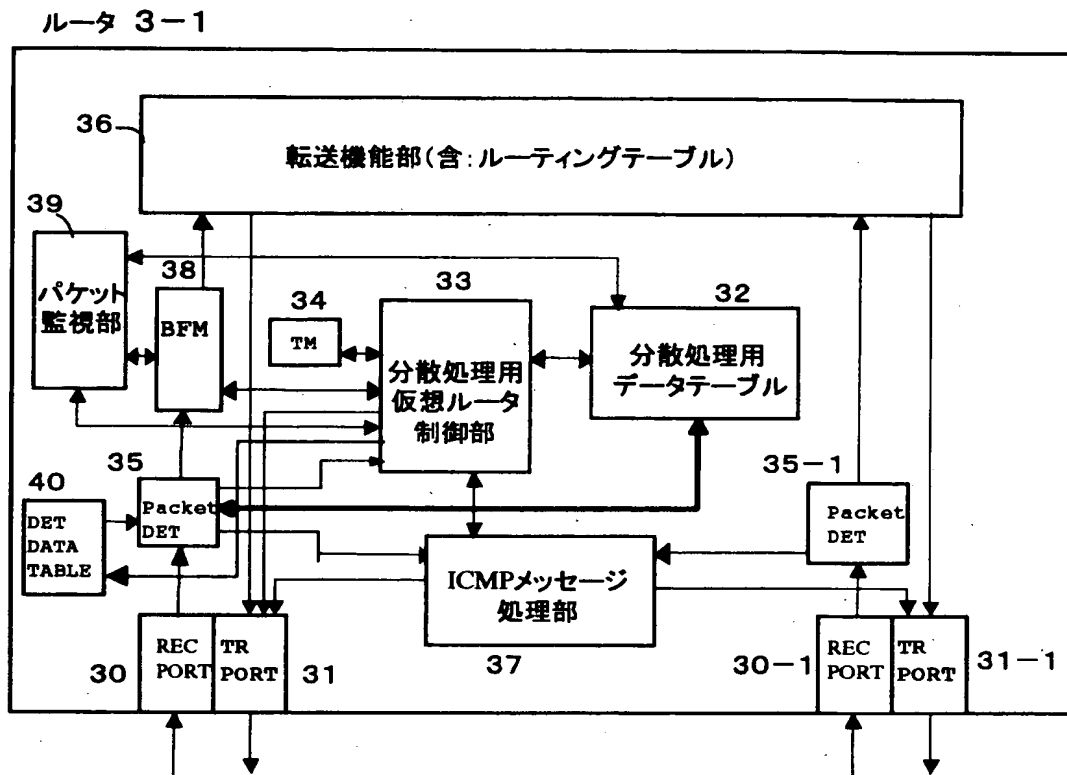
【図 2】



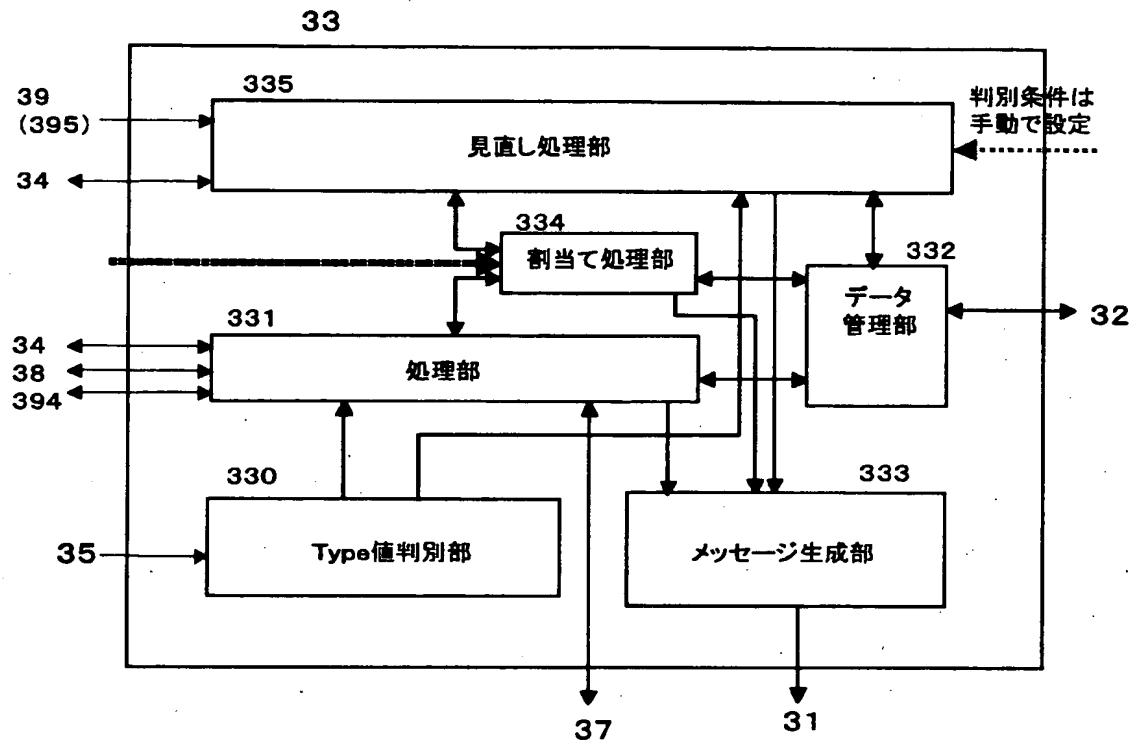
【図 3】

0		15,16		31	
Version	Type	VR ID	Priority	Count IP Address	
Auth Type		Adver Int	Check Sum		
IP Address (1)					
IP Address (2)					
⋮					
IP Address (n)					
Authentication Data (1)					
Authentication Data (2)					

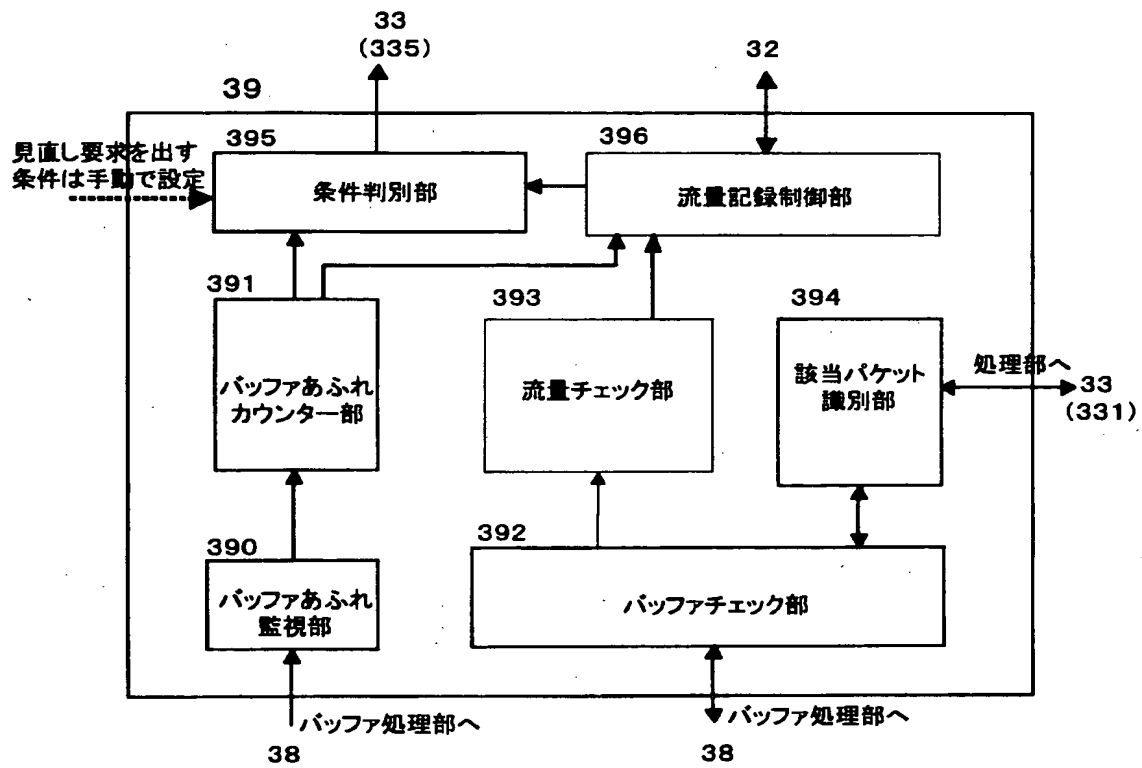
【図 4】



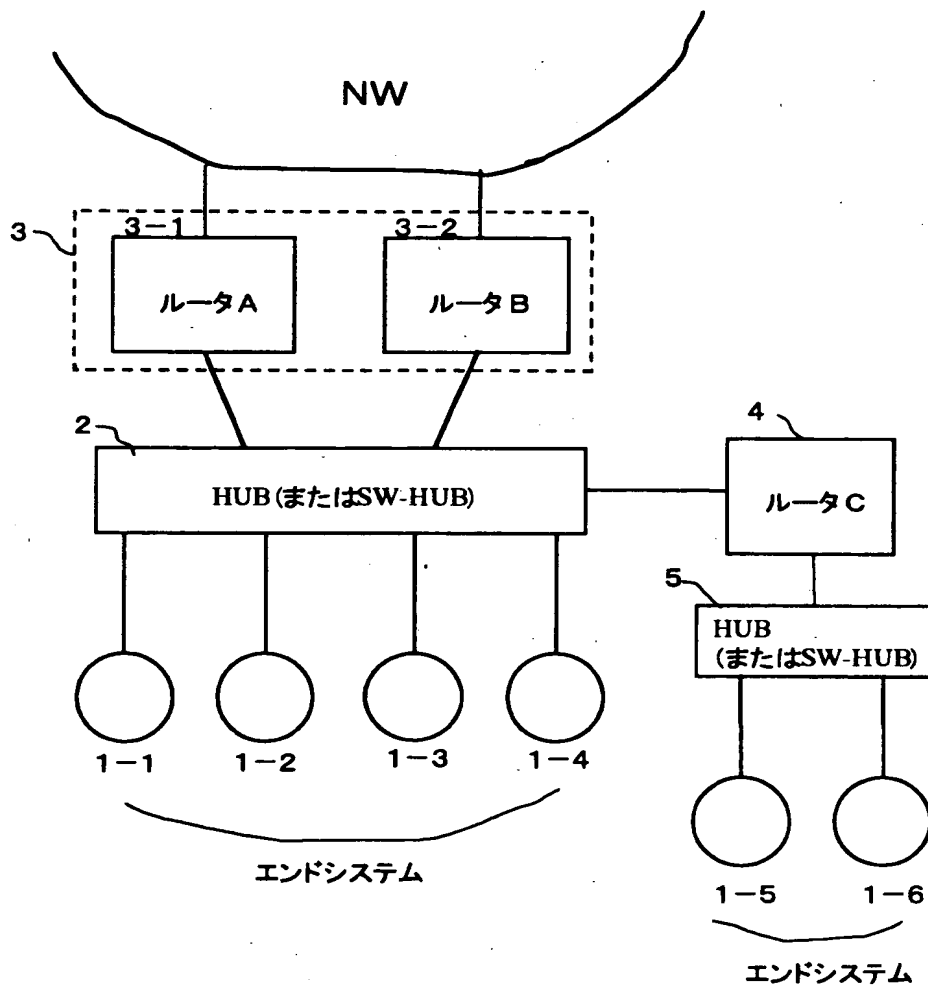
【図 5】



【図 6】



【図 7】



【図 8】

A

0		15,16		31	
バージョン	ヘッダ長	サービスタイプ	トータル長		
識別子 (ID)			フラグ	フラグメント・オフセット	
生存時間 (TTL)		プロトコル	ヘッダチェックサム		
送信元アドレス					
宛先アドレス					
オプション				パディング	

B

0		15,16		31
タイプ	コード	Check Sum		
識別子		順序番号		

【図 9】

分散処理用データテーブル

A 仮想ルータ情報

仮想ルータ ID	1
仮想IPアドレス	10. 15. 18. 1
動作モード	00
ルータ構成数	2
ルータ1	10. 15. 18. 1
ルータ1状態	0:稼働
ルータ2	10. 15. 18. 2
ルータ2状態	1:休止

B サブネット構成情報

ノード1	10. 15. 18. 100
ノード2	10. 15. 18. 101
ノード3	10. 15. 18. 102
⋮	⋮

C パケット情報

担当ルータのIPアドレス	条件(IPアドレス)	流量	バッファあふれ回数
10. 15. 18. 1	10. 15. 18. 100	XXXX	XX
	10. 15. 18. 101	XXXX	
10. 15. 18. 2	10. 15. 18. 102	XXXX	XX
⋮	⋮	⋮	⋮

【図 10】

拡張フレームフォーマット

A 割当てパケット

0	15,16				31
Version	Type	VR ID	Priority	Count	IP Address
Auth Type		Reserve	Check Sum		
ルータ1のIPアドレス					
対象のとなる装置数					
対象のIPアドレス1					
:					
ルータ2のIPアドレス					
対象のとなる装置数					
対象のIPアドレス1					
:					
Authentication Data (1)					
Authentication Data (2)					

※Type値は3

B 割当て確認パケット

0		15,16		31	
Version	Type	VR ID	Priority	Count IP Address	
Reserve			Check Sum		
ルータのIPアドレス					

※Type値は4

【図 11】

C シーケンス番号通知パケット

0	15,16				31
Version	Type	VR ID	Priority	Count IP Address	
Auth Type		Reserve	Check Sum		
宛先IPアドレス					
送信元IPアドレス					
TCPポート番号			TCPシーケンス番号		
:					
宛先IPアドレス					
送信元IPアドレス					
TCPポート番号			TCPシーケンス番号		
Authentication Data (1)					
Authentication Data (2)					

※Type値は5

D 見直し要求パケット

0	15,16			31
Version	Type	VR ID	Priority	Count IP Address
Reserve			Check Sum	

※Type値は6

E 流量通知要求パケット

0	15,16			31
Version	Type	VR ID	Priority	Count IP Address
Reserve			Check Sum	

※Type値は7

【図 1 2】

F 流量通知応答パケット

0		15.16		31
Version	Type	VR ID	Priority	Count IP Address
Auth Type		バッファあふれ回数	Check Sum	
対象の条件数				
対象のIPアドレス1				
流量				
⋮				
対象のIPアドレスn				
流量				
Authentication Data (1)				
Authentication Data (2)				

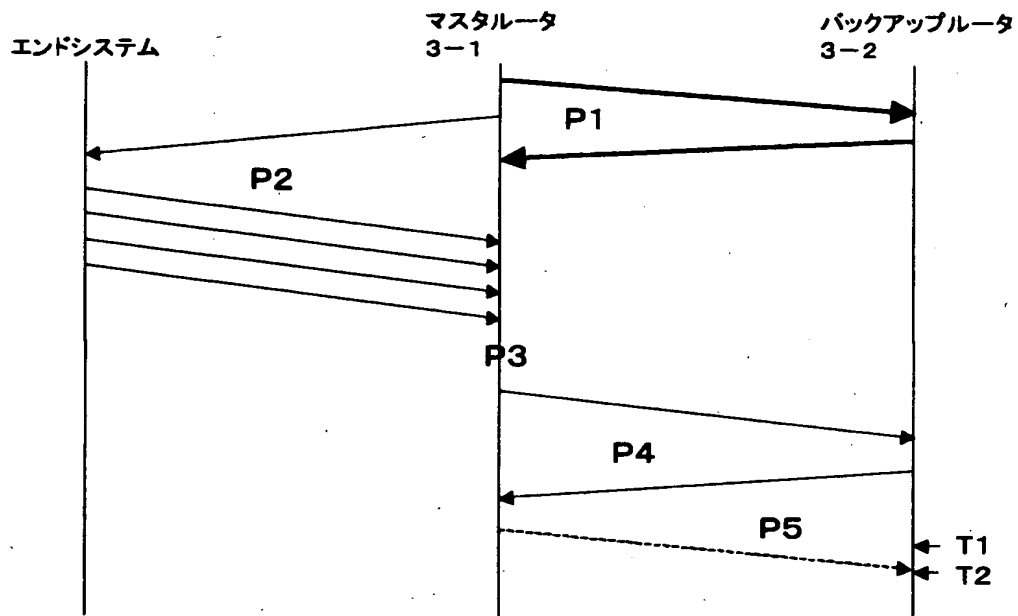
※Type値は8

【図 1 3】

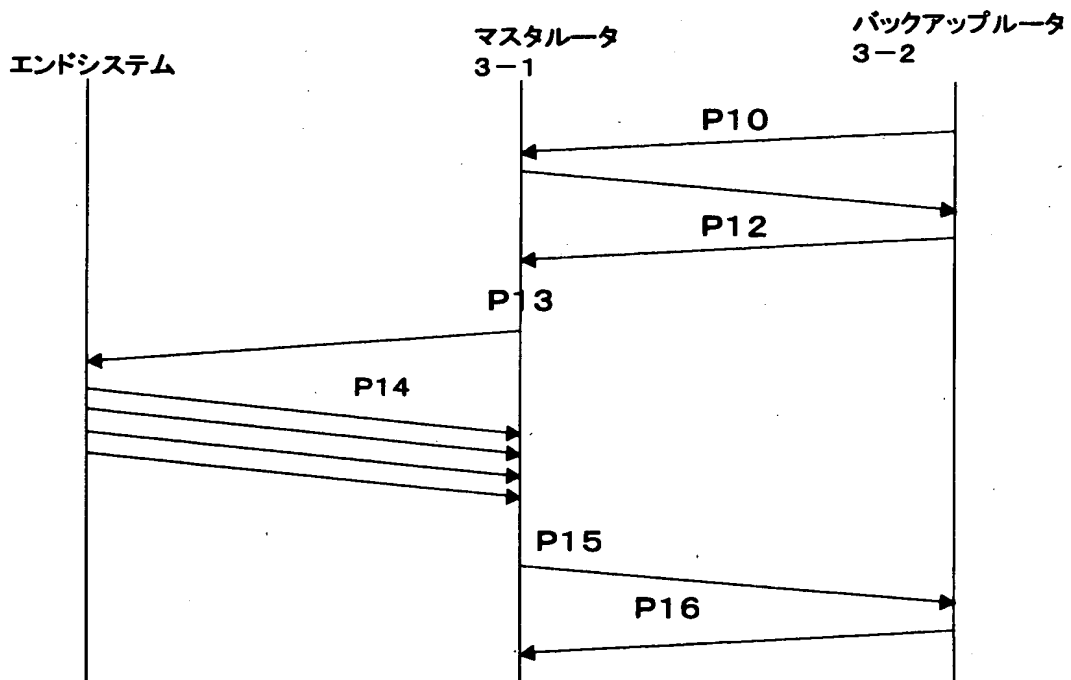
判別用データテーブルの一例

条件	状況
10. 15. 18. 100	通過
10. 15. 18. 101	破棄
10. 15. 18. 102	破棄
⋮	⋮

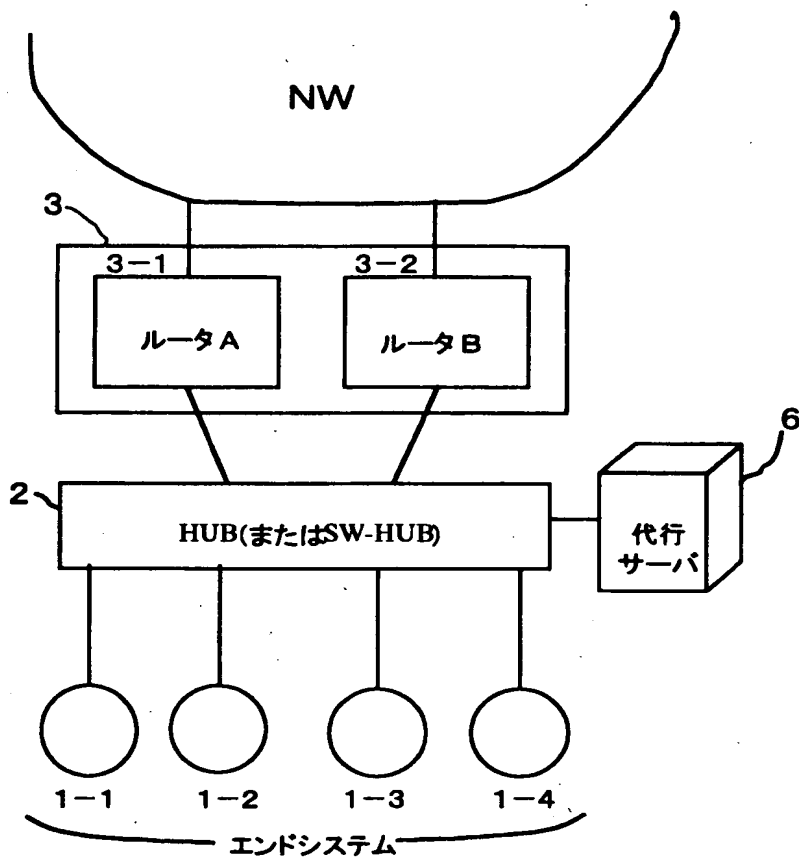
【図 14】



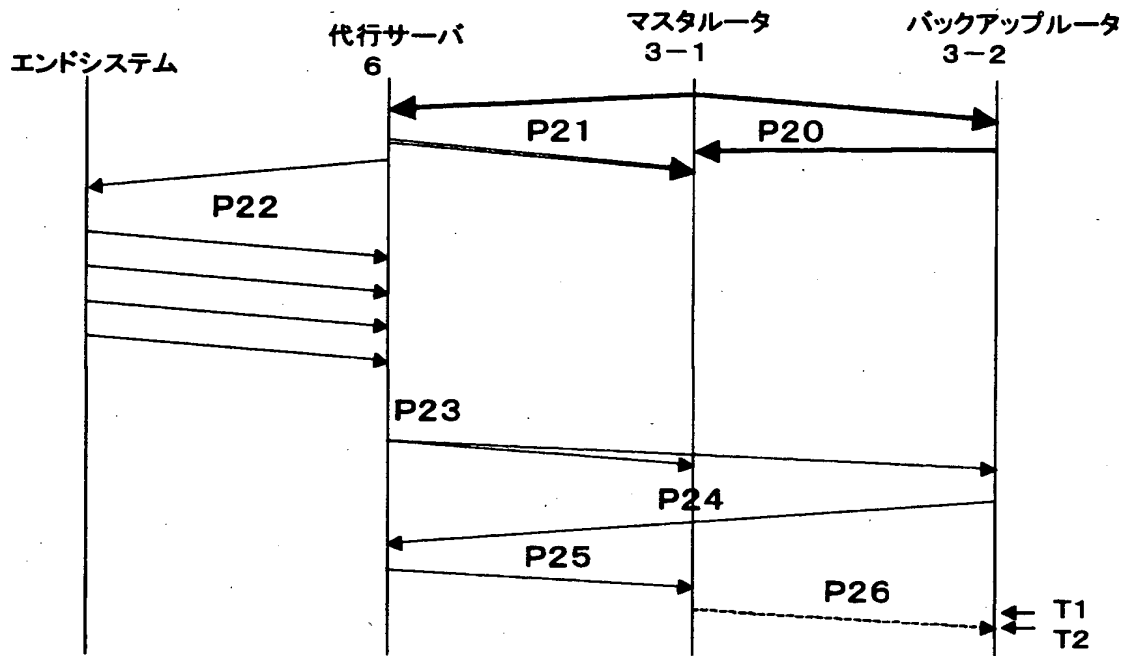
【図 15】



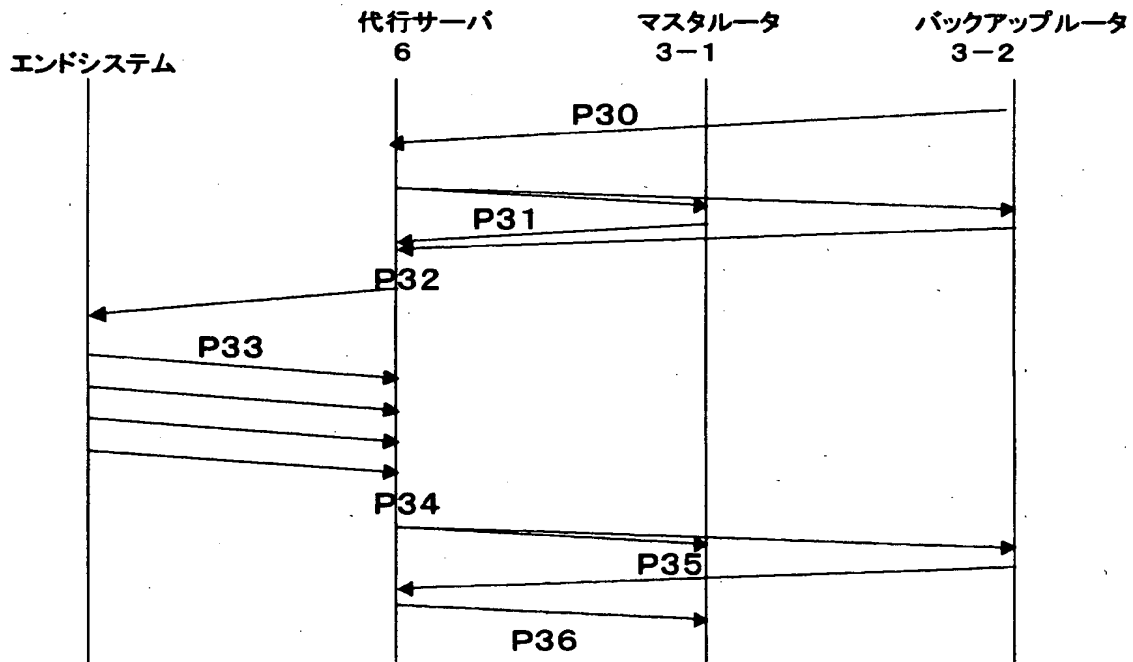
【図16】



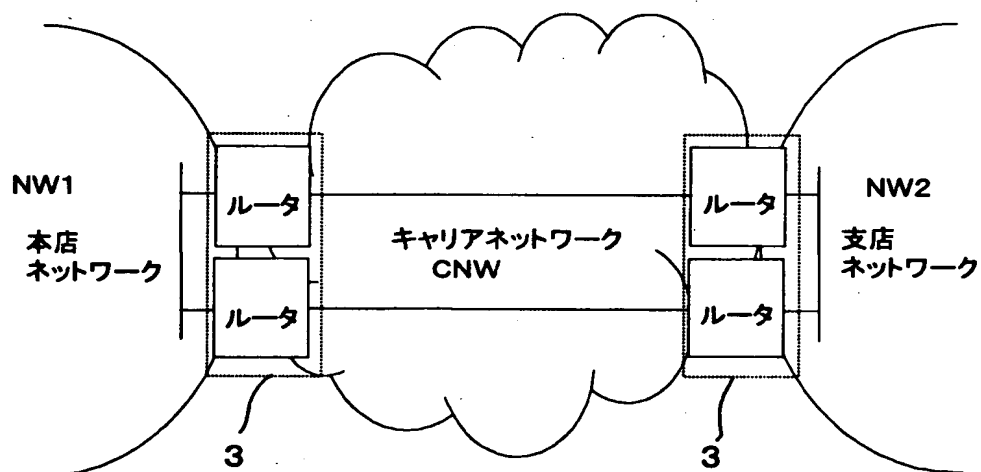
【図 17】



【図 18】



【図 19】



【書類名】 要約書

【要約】

【課題】 動的に負荷分散を実現する仮想ルータを利用した動的な負荷分散システムを提供する。

【解決手段】 1つの共通のアドレスを有する仮想ルータを構成する複数のルータ機能を有する装置と、前記仮想ルータを通して、ネットワークに接続されるエンドシステムを有し、前記仮想ルータを構成する複数のルータ機能を有する装置のうちのルータ機能を有する装置がマスタールータに設定され、他のルータ機能を有する装置がバックアップルータに設定され、前記設定されたマスタールータは、ルーティング対象となるパケットの条件を動的に設定し、該設定されたルーティング対象となるパケットの条件を前記バックアップルータに通知し、前記複数のルータ機能を有する装置により前記ネットワークとエンドシステム間のルーティング処理を行う。

【選択図】 図 1 4

認定・付加情報

特許出願の番号	特願2001-206856
受付番号	50100997717
書類名	特許願
担当官	佐藤 一博 1909
作成日	平成13年 7月13日

<認定情報・付加情報>

【特許出願人】

【識別番号】	000005223
【住所又は居所】	神奈川県川崎市中原区上小田中4丁目1番1号
【氏名又は名称】	富士通株式会社

【代理人】

申請人

【識別番号】	100094514
【住所又は居所】	神奈川県横浜市港北区新横浜3-9-5 第三東 昇ビル3階 林・土井 国際特許事務所
【氏名又は名称】	林 恒徳

【代理人】

【識別番号】	100094525
【住所又は居所】	神奈川県横浜市港北区新横浜3-9-5 第三東 昇ビル3階 林・土井 国際特許事務所
【氏名又は名称】	土井 健二

出 願 人 履 歴 情 報

識別番号 [000005223]

1. 変更年月日 1996年 3月26日

[変更理由] 住所変更

住 所 神奈川県川崎市中原区上小田中4丁目1番1号
氏 名 富士通株式会社